The Ohio State Technology Law Journal

TAX THEORY & FERAL AI

STEPHANIE HOFFER*

CONTENTS

* Professor of Law, The Ohio State University Moritz College of Law. This work benefited from the insightful comments of Ari Glogower, M. Shane Judy, James Kenney-Prentiss, participants in the Ohio State Technology Law Journal's 2019 symposium, Artificial Intelligence and the Future of Tax Law, participants in the 2019 annual conference for Mid-Career Tax Professors (AMT), and the capable work of research assistant Paul Mate and law librarian Susan Azyndar. Special thanks are owed to Mirit Eyal Cohen for her thoughtful AMT conference response and to Shu-Yi Oei for her encouragement. All mistakes are my own.

I. Introduction

This essay is a sci-fi thought experiment about the significance of personhood in income taxation, meant to explore the validity of currently prevailing justifications for the tax. Assume that the year is 2050. Developers, human or otherwise,[1] have created non-sentient artificial intelligences (AIs) capable of transacting in digital currency. Assume, perhaps improbably, that some of these AIs are "feral." A non-sentient AI might be feral in the future because it was never the property of a human, because it was abandoned by a human, or because it "escaped" into the wild. Imagine that non-sentient feral AIs create new value in the economy by doing things like writing, designing, securities and currency trading, planning, and 3D printing. They monetize that new economic value as active participants in the economy who sell goods or services to consumers. Intuition suggests that the value newly created by these independent economic actors should be included in the tax base and, in particular, the income tax base. Under current law, it is not.

The federal income tax law, and the theories that underpin it, have yet to fully address the status of non-human earners. Scholarship on AI and taxation primarily has focused on the taxation of AI's owners, on whether AI itself should be taxed despite being owned by someone else, or on the philosophical question of taxing *sentient* AI.[2] But

---

[1] The first patent applications for AI-made inventions were filed recently by an international team of researchers. *See* Lara Butler, *World first patent applications filed for inventions generated solely by artificial intelligence* (Aug. 1, 2019), https://www.surrey.ac.uk/news/world-first-patent-applications-filed-inventions-generated-solely-artificial-intelligence.

[2] *See generally* XAVIER OBERSON, TAXING ROBOTS: HELPING THE ECONOMY TO ADAPT TO THE USE OF ARTIFICIAL INTELLIGENCE (2019) (examining how current tax regimes would apply to robots owned by persons and considering whether robots could be granted legal personhood); Orly Mazur, *Taxing the Robots*, 46 PEPP L. REV. 277 (2019) (proposing rebalance of the tax burdens on labor and capital to adjust for potential decline in labor as a result of automation); Ryan Abbott & Brett Bogenschneider, *Should Robots Pay Taxes? Tax Policy in the Age of Automation*, 12 HARVARD L. & PUB. POLICY 145 (2018) (proposing countervailing tax measures such as disallowance of corporate tax deductions and introduction of an automation tax to remove tax incentives for automation of labor); Filipe Maia Alexandere, The Legal Status of Artificially Intelligent Robots: Personhood, Taxation and Control, 33 – 39 (June 1, 2017), https://dx.doi.org/10.2139/ssrn.2985466 (discussing tax proposals to mitigate the effect of automation of labor); Joachim Englisch, *Digitalisation and the Future of National Tax Systems: Taxing Robots?* (Sept. 5, 2018), https://dx.doi.org/10.2139/ssrn.3244670 (arguing

thinking about tax in the context of *non-sentient* feral AI allows us to identify the human elements and biases inherent in theoretical justifications of the income tax. Identifying those elements allows us to consider whether they are features or bugs. Specifically, this essay uses non-sentient feral AI as a vehicle to explore the application of two popular theoretical justifications for the income tax—the benefit principle and welfarist theories—and it concludes that a third, the ability to pay, is a better answer. The goal of this essay is not to suggest that feral AI should be subject to the income tax. Rather, it asks whether re-examining popular refrains in the tax policy canon within the context of feral AI will reveal anything about the human-centricity of those ideas and their application to humans.

Part II of the essay describes the current state of the law by walking through a series of hypotheticals involving machine earners, exploring the likely federal income tax consequences of each, to build a pragmatic (yet hypothetical) case for the inclusion of feral AI's earnings in the tax base. Part III of the essay addresses the objection that income taxes would not be necessary in the context of feral AI by exploring the application of natural resources or sales taxes to transactions between humans and non-sentient feral AI. It concludes that natural resource taxes and sales taxes would be inadequate substitutes for an income tax because they would leave AI undertaxed relative to its human counterparts in the income tax base. Part IV of the essay then asks whether income taxation of non-sentient feral AI's earnings would be supported by either the benefit principle or welfarist theories. Because those conceptions of taxation rely so heavily on human preferences, they are ill-fitting in the case of non-sentient AI. By analogy, they are also ill-fitting in cases when human preferences are weak, such as disposition of the last-earned dollars of the ultra-wealthy, or when human preferences are delegitimized by the government, such as those of marginalized groups like undocumented immigrants. Adopting a non-human point of view in tax policy, then, can inform our thinking about taxation outside of what humans perceive to be the norm.

---

that proposals to tax robots are premature. But taxation of the use of robots might slow disruption in the labor market).

II. Income Tax Law is Human-centric

Income tax law is, naturally, human-centric. It assumes that income is earned by people. An AI is not a person.[3] Without a "person" in the mix, earnings are not "income" under the Internal Revenue Code, and they are not subject to income taxation.[4] For tax purposes, a "person" is either a human individual or an entity that is owned directly or indirectly by human individuals.[5] A non-person, such as an AI or an animal with earnings potential, is excluded from income taxation, even when that non-person is not the property of an actual person. This essay explores why prevailing justifications for the income tax either support the exclusion the income of non-persons from the tax base or fail to consider it. The essay also will suggest that examining the gaps in tax policy surrounding non-persons can be instructive with regard the application of policy to persons.

To demonstrate the existing gap in tax policy with regard to non-owned non-persons like feral AI, the following paragraphs walk through three hypotheticals involving income earned by machines. In each case, current law dictates that the machine is not a taxpayer. In the first two examples, income earned by the machine instead accrues to whomever eventually exerts dominion and control over it. A final example discusses a hypothetical non-sentient feral AI not subject to capture by a person as a means of examining the roles of personhood and property rights in the current income tax regime. The section concludes that a combination of lack of property rights and lack of personhood creates a bar to currently including income of a non-human earner in the tax base. This section's discussion of current law serves as a prelude to Section III, which discusses how prevailing justifications of income taxation might apply in the case of non-owned non-persons like feral AI.

---

[3] As explained below, AI is excluded from the Internal Revenue Code's definition of person. Although proposals have been forwarded by some European politicians to grant legal personhood to robots, none have succeeded. *See* OBERSON, *supra* note 2, at 131.

[4] *See, e.g.*, I.R.C. § 1 (2019) (income tax imposed on individuals); I.R.C. § 11 (2019) (income tax imposed on corporations); I.R.C. § 641 (2019) (income tax imposed on trusts and estates).

[5] I.R.C. §7701(a)(1) (2019).

a. Example 1

Here is the first example. Imagine an artist who creates a non-sentient robot that rolls along the sidewalk collecting lost coins. The artist, for the purpose of performance art, abandons the robot. Ownerless, it wanders around picking up pennies, nickels, dimes, and quarters. If it were a person, it would be taxed on the money it collected. The loose change would be a treasure trove.[6] The robot would have "an undeniable accession to wealth, clearly realized, and over which the taxpayer has complete dominion," which is the commonly accepted definition of income subject to taxation under federal law.[7] But the robot is not a person. Should it be taxed on its income? The conventional answer is no. It is nothing more than abandoned property, and it has no property rights in anything that it possesses. The robot itself, either empty or full of coins, is a treasure trove to be included in the income of the person that eventually exerts dominion over it.[8]

b. Example 2

Consider a second example. Imagine that the artist creates a second non-sentient robot. This new robot rolls along the sidewalk collecting discarded objects, which it uses to make found-object art. Again, the artist abandons the robot. Now imagine that the robot is a mobile vending machine. It can accept money in exchange for art. This scenario is slightly different from the first. Unlike the first robot, the second robot uses labor to take something from the commons and create new economic value where none existed before.[9] Also unlike the first robot, the second robot engages in exchanges for value. It both expands the economy and has realization events that would trigger taxation for a person. But the robot is not a person. Should the second robot be taxed on its income? Again, the conventional answer is no,

---

[6] Cesarini v. United States, 296 F. Supp. 3, 7 - 8 (1969) (found property or money is a "treasure trove" taxable to the finder whose possession is undisputed).
[7] Charley v. C.I.R., 91 F.3d 72, 74 (9th Cir. 1996).
[8] *Cesarini*, 296 F. Supp. at 7 – 8.
[9] John Locke might have advocated for property rights on this basis had the actor had been human. *See* JOHN LOCKE, SECOND TREATISE OF GOVERNMENT (C.B. Macpherson ed., Hackett Publishing Co. 1980)(1690).

for all of the reasons given above.[10] This second scenario is more interesting, though, because the robot has taken on some characteristics that traditionally are associated with taxpayers. It is generative, and, until someone seizes it, it exerts control over the value it generates. The problem remains, though, that like the first robot, the second robot is abandoned property and will be a treasure trove taxable to the person that eventually exerts dominion over it.[11]

c. Example 3

Here is a third and final example, from which the rest of the essay proceeds. In it, the artist creates a "robot" that is in the digital world rather than in the purely physical one. Imagine that the artist creates a non-sentient AI and releases it into the digital world, abandoning any means of reclaiming it. Assume that, like the second robot, the AI creates value in the economy and engages in exchanges. For instance, it may generate written content, trade digital goods, or make art. Also assume that the AI uses cryptocurrency. Finally, assume that exerting dominion over the AI would be difficult or impossible.[12] Should the AI be taxed on its income, ignoring for the moment the practical impossibility of doing so?[13]

---

[10] Although notice that a VAT could easily apply in this context. *See* OBERSON, *supra* note 2, at 87 – 110 (suggesting ways in which a value added tax might apply to AI or robots).

[11] *Cesarini*, 296 F. Supp. at 7 – 8.

[12] Cryptocurrencies are able to evade government regulation due to the use of blockchain technology that makes transactions with such currencies nearly impossible to track. To illustrate this point, Canadian entrepreneur and founder of Quadriga, Canada's largest cryptocurrency exchange, died early in 2019, leaving behind $145 million dollars in Bitcoin and other cryptocurrency assets that are currently inaccessible to the cryptocurrency's investors. Daniel Shane, *A crypto exchange may have lost $145 million after its CEO suddenly died*, CNN Business (Feb. 5, 2019), https://www.cnn.com/2019/02/05/tech/quadriga-gerald-cotten-cryptocurrency/index.html. The founder's widow, Jennifer Robertson, said that the laptop used by her late husband to run the currency exchange is encrypted, and she does not know the password or recovery key. *Id.*

[13] Note that absent some extraordinary programming, if a government were able to enforce collections against the AI, it is likely that a person could exert dominion over the AI, which would bring our analysis back to treasure trove. Because this essay is a though experiment meant to ask what the impossible can teach us about the possible, let's set the question of collections aside.

Whether feral AI should owe income tax on its digital earnings is not easily resolved.[14] The usual "no" argument remains that property cannot own property, cannot have wealth, and therefore cannot have an accession to wealth. Technically, then, it fails to meet the federal definition of income. On the other hand, ownership is neither a necessary nor a sufficient condition for becoming a taxpayer. For instance, embezzled funds are treated as income to the embezzler, even though someone else owns them.[15] Assigned income is treated as income of the assignor even if it legally is property of the assignee.[16] Partnerships are not taxpayers even though they legally own income that they earn; rather, the income is reported by the partners.[17] Transfers of payments to AI are conceptually no different from other transfers. The determination of the appropriate taxpayer for AI-counterparty transactions should hinge on who has dominion and control over the transferred property—so, the AI. Or, because feral AI has no legal right to interact with a counterparty, perhaps such transfers are analogous to abandonment by the human counterparty followed by dominion asserted by the AI. Regardless, reliance on AI's lack of legally-enforceable property rights provides an insufficient justification for non-taxation. The income tax law is more concerned with substantive dominion and control than it is with formal property rights.

The prior paragraph established that lack of property rights should not be a bar to taxing non-sentient feral AI, but what about the AI's lack of personhood? Under current law, non-sentient feral AI is not a person for purposes of the Internal Revenue Code. Section 7701 provides that "[t]he term 'person' shall be construed to mean and include an individual, a trust, estate, partnership, association, company, or corporation."[18] AI, absent human intervention, is not a trust, estate,

---

[14] Notice that the question is not whether feral AI *could* be taxed. If we infer that it could be taxed, its assets would be seizable, and if its assets are seizable, it is nothing more than a treasure trove.

[15] *See* James v. United States, 81 S. Ct. 1052, 1055 (1961) (illegal gains included in gross income).

[16] *See* Lucas v. Earl, 50 S. Ct. 241 (1930) (earnings must be taxed to the earner even if a third party has a legal right to them once they are earned).

[17] I.R.C. § 701 (2019).

[18] I.R.C. §7701(a)(1) (2019).

partnership, association, company, or corporation (although if not feral it could be owned by one of those). Consequently, for feral AI to be subject to tax under current law, it must be an "individual."

Section 1 of the Code imposes an income tax on individuals.[19] Although the word "individual" appears with regularity in the Code, it is not defined. The presumption seems to be that an individual is a human person. Instances where the word "individual" includes non-humans are specifically noted, and none of those instances (of course) refers to AI.[20] By contrast, "individuals" in the Code do lots of things that currently only humans do, like get married,[21] buy a home,[22] have children,[23] go to jail,[24] and retire.[25] It seems unlikely that non-sentient feral AI would be an "individual" for purposes of the current income tax, since it neither engages in typical human activities nor is an entity statutorily designated as an individual.

In conclusion, under current federal income tax law, non-sentient feral AI is not a proper object of taxation because it is neither an individual nor a legal person. But why is personhood relevant to the inclusion of income in the tax base? Personhood connotes the right of ownership, but as noted earlier, ownership is not required for income taxation. Is personhood merely a signal for other, more normatively relevant characteristics, like the ability to create new economic value, remove and segregate value from the economy, and engage independently in economic exchange? The hypothetical above assumed that non-sentient feral AI would be capable of those things. If so, by analogy,

---

[19] I.R.C. § 1 (2019). In addition, section 11 imposes income tax on corporations, but our hypothetical AI is not a corporation.

[20] For example, *see* I.R.C. § 542(a)(2) (2019) ("For purposes of this paragraph, an organization described in section 401(a), 501(c)(17), or 509(a) or a portion of a trust permanently set aside or to be used exclusively for the purposes described in section 642(c) or a corresponding provision of a prior income tax law shall be considered an individual") and § 165(l)(2) (2019) (defining "qualified individual" as including related persons under IRC § 267(b), which includes corporations).

[21] I.R.C. § 7703 (2019).

[22] I.R.C. § 36 (2019).

[23] I.R.C. § 151 (2019).

[24] I.R.C. § 139F (2019).

[25] I.R.C. § 408 (2019).

the hypothetical AI should be taxed whether or not it has legal personhood.

From a position of pragmatism, new value in the economy should be included in the tax base, even when the value is created by a non-owned, non-human actor like a feral AI. It is not obvious, however, that income taxation is the appropriate mechanism for inclusion. This essay focuses on income taxes, but sales taxes or taxes on the extraction of natural resources are other potential ways of including newly added value in the tax base. The next section discusses those tax regimes and concludes that for non-sentient feral AI, natural resource taxes are an uneasy fit and that the application of sales or value-added taxes does not preclude the use of an income tax.

III. Natural Resource Taxes, Sales Taxes, and Feral AI

In both the federal and state systems, the income tax is joined by a variety of other taxes, including sales taxes and natural resources taxes. Although they are not relevant to this essay's exploration of benefit and welfarist theories of taxation below, I include them as a means of highlighting the importance of the income tax as a driver of distribution of the tax burden.

a. Sales Taxes

An objection to exploring income tax theories in the context of non-sentient feral AI may be that other taxes, such as natural resource or sales taxes, will suffice, so that the income tax may be reserved for people without compromising the integrity of the tax base. Although the goal of this essay is not to champion the taxation of feral AI, but rather is to explore what an AI-centered thought experiment can teach us about tax policy as it relates to humans, this section addresses the objection that income tax would not be needed in the case of non-sentient feral AI.

One method of taxing feral AI may be to subject its transactions to sales tax. Sales taxes typically are charged against the sales of goods

and services to end consumers.[26] Although there is no federal sales tax yet, they are commonly levied by the states.[27] The tax typically takes one of three forms.[28] A consumer levy is charged against customers but is collected by the seller, and it is the most common form of sales tax.[29] A privilege tax falls upon the seller for the privilege of doing business in the state, but the incidence of this tax probably is shifted to the customer through pricing.[30] Finally, a transaction tax is charged against the transaction itself, with liability for payment falling on the customer and liability for collection falling on the seller.[31] In the case of feral AI, it is tempting to think that sales taxes in any of these forms might be sufficient to reach earnings. A sales tax is particularly attractive because, unlike the income tax, it does not turn on the AI's personhood or property rights. Instead, it is triggered by a transaction.[32]

A handful of problems prevent sales taxes from effectively reaching the hypothetical earnings of our imaginary AI. First, consumers are liable for sales taxes in most jurisdictions. Non-sentient feral AI would be a seller rather than a consumer, so a sales tax would leave the earnings of the AI unaffected in consumer-based jurisdictions.[33] In other words, even if transactions with the AI were subject to sales taxes, feral AI would remain undertaxed in comparison to human or corporate merchants in consumer levy states and transaction tax states first because the AI would never be a consumer and second because other merchants would be subject to both sales and income taxes. This would be true even if consumer levy states and transaction tax states enacted a privilege tax specifically on feral AI. Doing so would create parity for purposes of the sales tax but not the income tax. It also would be true in privilege tax states themselves, because once again the AI would not be liable for income taxes that other sellers must pay. A second reason why sales tax on AI counterparty transactions is an

---

[26] 1 BENDER'S STATE TAXATION: PRINCIPLES AND PRACTICE § 12.01 (2019).
[27] *Id.*
[28] *Id.*
[29] *Id.*
[30] *Id.*
[31] *Id.*
[32] *Id.*
[33] *Id.* This assumes that the AI's pricing is inelastic.

insufficient stand-in for the income tax is that although the consumer is nominally liable for sales taxes in most states, the merchant is the designated collection agent for the tax.[34] Feral AI may not have been programmed to collect and remit sales taxes and, even if it were (since we are dealing in fictions already), it likely would not be able to respond adequately to changes in the law. Third, there is no federal sales tax, so absent a sea change in federal tax policy, the federal government could not use sales taxes to reach feral AI even if the incidence of tax were not in question.

Sales taxes seemingly are not a good solution to the income tax problem.[35] Even if sales taxes applied, this essay's primary inquiry– the question of philosophical support for an income tax on non-sentient feral AI— would remain unanswered.

b. Natural Resources Taxes

If feral AI is not a person for tax purposes, could it be a natural resource? Or could we view the data that it uses as a natural resource and its human customers as extractors?

Feral AI, like a feral animal, could conceivably be categorized as a resource.[36] If captured, it would be a treasure trove to the person who exerts dominion over it,[37] or perhaps capturing it would result in untaxed imputed income, like fish to a fisherman or a feral hog to a hunter. And if a person could assert dominion over feral AI, found property and captured animals would be apt analogies. For purposes of philosophical exploration, though, let us continue to hypothesize that it would be difficult or impossible to subject feral AI to the dominion of a person. Also recall that unlike feral animals who produce additional economic value (for instance, by giving birth to other feral animals of

---

[34] This is true regardless of the form that the tax takes. *Id.*

[35] *Cf.* OBERSON, *supra* note 2, at 87 – 110 (suggesting ways in which a value added tax might apply to AI or robots).

[36] Many thanks to the annual conference of Mid-Career Tax Professors, and especially Miranda Perry, for a completely entertaining discussion of this idea. Notes on file with author.

[37] *Cesarini*, 296 F. Supp. 3 at 7 - 8 (found property or money is a "treasure trove" taxable to the finder whose possession is undisputed).

value), our hypothetical feral AI is capable of monetizing its newly created value by engaging in transactions with people.[38] Consequently, for purposes of this essay, it makes little sense to treat feral AI like a feral animal.[39]

Even if non-sentient feral AI differs meaningfully from a feral animal under the constraints of our hypothetical, the natural resource analogy still may apply to its customers. Perhaps a customer of AI in the digital world would be analogous to a miners who extract resources from the physical one. Instead of drilling for oil or mining for coal, AI would mine data.

In the physical world, resource extraction may be subject to severance taxes. Currently, thirty four states levy severance taxes on the extraction of intrastate natural resources.[40] For example, in Alaska, a severance tax is levied on the producer of oil or natural gas.[41] For oil, the tax is levied at "five percent of the gross value at the point of production of the oil," while for natural gas, the tax is levied at "1.667 percent of the gross value at the point of production of the gas."[42] Likewise, in Montana, a severance tax is levied on coal mined within the state.[43] Under Article IX, §5 of the Montana state constitution, the legislature is directed to contribute at least half of the coal severance

---

[38] *Cf* Inventor Trains Crows to Find Money, https://www.npr.org/templates/story/story.php?storyId=87878028 (March 4, 2008) (scientist trained crows to collect coins and exchange them for food). Consider an example suggested to at the Association of Mid-Career Tax Professors in 2019 by Michael Simkovic: a homeowner with a mouse problem leaves food out for a feral cat. In exchange, the cat enters the home and kills mice. Does the cat have income equal to the market value of the food that it consumes? Does the homeowner have income equal to the market value of the extermination of the mice? *See* Rev. Rul. 79-24, 1979-1 C.B. 60 (recipients of goods or services in barter transactions have gross income equal to the value of what they received).

[39] I hope to relax this essay's assumptions and develop the feral animal comparison more thoroughly in future work on the taxation of real, rather than hypothetical, AI.

[40] The State of State (and Local) Tax Policy, Tax Policy Center, Urban Center and Brookings Institute, https://www.taxpolicycenter.org/briefing-book/how-do-state-and-local-severance-taxes-work (last visited on August 30, 2019).

[41] Oil and gas production tax, Alaska Stat. Ann. § 43.55.011(e) (West 2019).

[42] *Id.*

[43] It is also important to note that Montana is also the state in which the U.S. Supreme Court ruled that the intrastate collection of a severance tax was ruled to be constitutional, because such a tax does not interfere with the Commerce Clause or the Supremacy Clause. *Commonwealth Edison Co. v. Montana,* 453 U.S. 609 (1981).

tax collected to a trust fund, and "the principle of the trust shall remain inviolate unless appropriated by vote of three-fourths (3/4) of the members of each house of the legislature."[44] Through the trust fund, Montana issues bonds used to finance infrastructure projects and "renewable resource projects in the state to provide [...] a healthy economy, alleviation social and economic impacts created by coal development, and a clean and healthful environment for future generations."[45] Many other states have severances taxes as well.[46]

The comparison of transactions with non-sentient feral AI to natural resource extraction is attractive but not complete. On one hand, AI is similar to equipment used in mining. In fact, it literally may be mining data. And perhaps data is similar to a natural resource because much of it exists in a public commons. To the extent that an economic actor profits from extraction of this common resource, it may be equitable to exact a fee for the benefit of those who have an interest in the commons, just as Alaska and Montana do. On the other hand, this argument would apply not only to feral AI but also to AI that is owned by a taxpayer. In addition, extraction taxes typically apply to the extractor who, in this analogy, would be the user of AI rather than the AI itself. So like a consumer or transaction-based sales tax, an extraction tax on feral AI would not reach the earnings of the AI, but rather would increase the cost of the transaction to the user, assuming that the AI's pricing is inelastic. As with sales taxes, the AI would be relatively undertaxed in comparison to market actors who are persons subject to income taxes. Finally, unlike coal or helium or natural gas, data in a commons are usually non-rivalrous. While many natural resource taxes have as a goal the restoration of value to a jurisdiction's

[44] Mont. Const. art. IX, § 5.
[45] Mont. Code Ann. § 17-5-701 (West 2019).
[46] A few additional examples follow. In Michigan, a severance tax is levied on oil or gas. Mich. Comp. Laws Ann. § 205.301 (West 2019). In North Dakota, a severance tax is levied on "all coal severed for sale or industrial purposes by coal mines within the state" at $37.50 per ton of 2,000 pounds. N.D. Cent. Code § 57-61-01 (2019). According to the statute, the tax is "in lieu of any sales or use taxes imposed by law," and "[e]ach coal mine owner or operator shall remit the tax for each month, within twenty-five days after the end of each month, to the tax commissioner on reports and forms as the tax commissioner deems necessary." *Id.* In West Virginia, an annual minimum severance tax of $0.50 is imposed per ton on the "severing, extracting, reducing to possession or producing coal for sale, profit or commercial use." W.Va. Code, § 11-12B-3 (2019). Other examples can be found in state statutory compilations.

residents in light of depletion, the same reasoning likely would not apply to feral AI's use of data (although a compensatory justification would remain).

In summary, while sales taxes or extraction taxes on non-sentient feral AI may raise funds by increasing the cost to the AI's users, they would leave the AI's earnings untouched. As a result, the question of an income tax remains relevant.

The next section discusses whether currently prevailing theoretical justifications for the income tax undergird or work against the pragmatic case for taxing feral AI. Where gaps in the tax policy exist as it applies to non-sentient feral AI, the essay questions whether a human-centric view of tax policy hinders the development of tax policy not only in the sci-fi world but also in the real one beyond.

IV. Prevailing Theoretical Justifications for the Income Tax Are Human-Centric

A pragmatic case for taxing non-sentient feral AI is that, in the hypothetical above, it creates economic value that is excluded from the income tax base. By taking pre-existing information and adding value, AI increases the overall size of the economic pie without sharing a slice. Other creators of value—people and assets—are covered by provisions for the taxation of ordinary income and capital gains, but none of these provisions in their current iterations would cover feral AI. If, in the third hypothetical above, we ignore the AI because it is neither a person nor an asset owned by a person, newly created economic value will fall permanently outside of the tax base. As a practical matter, the AI's gains are analogous to the income of a person, and governments should want to capture some of this new value through income taxation. At the same time, many existing theoretical justifications of the income tax are person-centric, and AI is not a person (yet). The following paragraphs discuss two of those justifications, the benefit principle and generalized welfarist theories, as a way of exploring human-centric elements of the philosophy behind income taxation. Specifically, the satisfaction of human subjective preferences looms large in both the benefit principle and

welfarist theories, and it creates distortion in cases when humans have no preferences, are unable to express them, or have preferences that are either disregarded or non-cognizable by the government.

a. The Benefit Principle of Taxation

The benefit principle of taxation holds that people should bear the cost of government in proportion to the benefits that they receive from it.[47] In its simplest form, a benefit-based tax would be the equivalent of a user fee for something like postage or utilities, levied on the basis of a taxpayer's revealed preferences for and use of government-provided goods and services.[48] Traditionally, the idea's proponents posited that the "since social contract required the government to perform various tasks (such as protecting life or property), and individuals were obliged to fund those services, the extent that individuals benefited from those contracted services determined their liability to tax."[49] In other words, taxpayers would have a preference for the level of public goods provided by a government, would reveal that preference, and would pay accordingly.[50]

The benefit principle is problematic when applied to non-market goods such as the protection of rights, which are difficult to value, or when applied across a large number of people where not everyone reveals or receives his or her preferred level of public goods.[51] Modern scholarship has attempted to resolve these issues by using a taxpayer's income as a proxy for the benefit that the taxpayer derives from government, since government protection and infrastructure are required to create and preserve income.[52] If income is an adequate

---

[47] JOSEPH J. CORDES ET AL., THE ENCYCLOPEDIA OF TAXATION & TAX POLICY 24 (2d ed. 2005); JOEL SLEMROD & JON BAKIJA, TAXING OURSELVES 62 (4th ed. 2008).
[48] SLEMROD & BAKIJA, *supra* note 47, at 62.
[49] Graeme S. Cooper, *The Benefit Theory of Taxation*, 11 AUSTRL. TAX F. 397, 432 (1994).
[50] *Id.* at 448. *See also* RICHARD A. MUSGRAVE, THE THEORY OF PUBLIC FINANCE, A STUDY IN PUBLIC ECONOMY 69 (1959) ("Taxation according to benefits received was to be formulated by determining tax shares according to subjective evaluation of public services.").
[51] *Id.* at 452; SLEMROD & BAKIJA, *supra* note 47, at 62-63.
[52] Cooper, *supra* note 49, at 484, SLEMROD & BAKIJA, *supra* note 47, at 63.

proxy for benefit, the benefit principle may be used to justify progressive income taxation.[53]

In the context of income earned by feral AI, the benefit principle is less cogent, and examining it in that context brings to light built-in human assumptions and biases. As discussed above, satisfaction of preferences is central to the benefit principle, and preferences, as they are conceived of in political theory, are inherently subjective. For instance, Richard Musgrave wrote in 1959 that the benefit principle "has the great merit of tying the choice of public services to the preferences of individual members of the community."[54] Other scholars have noted that the benefit principle might "lead taxpayers to make better-informed decisions about the size and scope of government,"[55] since "[y]ou certainly can't just ask people what government activities like national defense are worth to them."[56] These comments and others like them demonstrate a focus on what taxpayers want, and they assume that the taxpayers in question can express at least some approval or disapproval through civic engagement.

The benefit principle's subjective focus on what people want makes little sense in the context of a hypothetical feral AI.[57] Non-sentient AI is incapable of having preferences. It cannot "experience" benefit. It may, for example, accrue more or fewer assets as a result of environmental factors; but it is neither satisfied nor dissatisfied as a result. To say that AI benefits from the existence of markets or the internet anthropomorphizes AI based on human preferences about trade. Non-sentient AI, being nothing more than a set of algorithms, is incapable of forming preferences about the existence of markets or the internet, so it is not advantaged or disadvantaged by any government act affecting those things. It might function more or less efficiently in response to a government act, but because it is not self-aware, it does

---

[54] MUSGRAVE, *supra* note 50, at 62.
[55] *See* CORDES ET AL., *supra* note 47, at 24.
[56] SLEMROD & BAKIJA, *supra* note 47, at 62.
[57] This is not to criticize Professor Oberson's conclusion that the benefit principle could apply in the context of AI owned by a person. *See* OBERSON, *supra* note 2, at 34. He notes that fees for supervision or registration of robots could fall under this umbrella but that they would not be a major contributor to tax revenue. *Id.* at 43.

not experience fluctuations in efficiency as beneficial or detrimental. To say that they *are* beneficial or detrimental because the AI engages in more or fewer transactions projects our own beliefs about worth onto an actor that is essentially just a mechanical process.

That non-sentient feral AI is unable to experience benefit should, under the benefit principle, lead to the conclusion that income earned by it falls outside of the tax base. And yet, new value created by feral AI does not differ substantively from new value created by any other taxpayer. As a result, the benefit principle, at least in its subjective formulation, falls short as a stand-alone means of distributing the income taxation burden in this circumstance.

The failure of the benefit principle in the case of feral AI highlights a very human failing of the idea itself. That subjective benefit is difficult to quantify has long been a shortcoming the benefit principle, particularly given the reach of the modern administrative state in daily life.[58] But the failing is more fundamental than simply the difficulty of appraisal. The idea itself—that the government, along with the rights it provides and protects, can be assigned a dollar value—is jaundiced. This is particularly true for that iteration of the benefit principle that uses income as a proxy for benefit.

In any of its iterations, the benefit principle leaves open the question of how to appraise the provision and protection of more ephemeral rights like voting, free speech, the use of public spaces, and privacy. When income is employed to measure benefit, we are forced to assume that rights not related to income and property, or other hard-to-value benefits, like the psychic benefits of government, are shared either equally by all (and therefore don't affect the distribution of tax burden) or are shared in proportion to income (and therefore are one of the drivers of distribution of the tax burden). To assume the former ignores political failings with regard to marginalized people or groups who may not have equal access to the rights and protections afforded to others by the government. To assume the latter suggests that the market value of a taxpayer's labor or return to capital determines, for

[58] *Id.* at 62 – 63.

appraisal purposes, the value of that taxpayer's rights to free speech, privacy, etc. If so, rights are inherently more valuable to high earners than to low earners. Such a view fails to afford equal dignity to taxpayers at the low end of the income spectrum, and it falls apart completely when one considers the juxtaposition of a human earner with non-sentient feral AI.

Substituting a non-sentient feral AI for a marginalized taxpayer or a low income taxpayer in the analysis above draws into sharp focus the failure of the benefit principle to make allowances to taxpayers who either cannot express preferences about government or whose preferences about government are marginalized. Two examples demonstrate the problem of using income as a proxy for benefit in this regard. First, consider the juxtaposition of a low-earning human and a high-earning feral AI. Under a version of the benefit theory that employs income as a proxy for benefit, the AI will have higher income tax liability than the human because the AI is deemed to have benefitted more from the existence of the government. But the human has rights while the AI does not. A conclusion that the AI benefits more from government than the human discounts the value of those rights. After all, the AI cannot prize more highly than a human being rights that it does not possess. The message sent by the benefit principle here is that ephemeral rights have little or no real value.

Dismissing income as a proxy for benefit and choosing subjective valuation instead produces a different undesirable result. Consider as a second example two market participants: one is a feral AI, and the other is a human being. Both earn $50,000 during the year writing advertising content. From a subjective standpoint, the AI values the existence of the government at $0, since it can neither have preferences nor experience the satisfaction of them. The human taxpayer, on the other hand, likely values the existence of government at more than $0. If the taxpayer's subjective valuation of government benefits is used to determine tax liability, the human taxpayer will have greater liability than the AI. In sheer dollar terms, the result is inequitable. Both earners have created $50,000 of economic value, but their tax burden differs. A higher tax burden will put the human at a competitive disadvantage vis-à-vis the AI. Thinking about AI in this

context highlights the ways in which human-centricity compromises the benefit principle as a foundation of tax policy.

b. The Principle of Welfare Maximization

Welfarist theories of taxation are another predominant strand of thinking that may be troubled in the context of non-sentient feral AI. They gauge distributive justice based on the social welfare or utility of people within a polity.[59] To analyze a tax system through a welfarist lense, policy theorists first determine the social welfare or utility of individuals and then aggregate those individual amounts.[60] A system that produces more aggregate utility or welfare is better than one that produces less. Optimal tax theory, a more robust version of this approach, balances both welfare and economic efficiency.[61] Although there is no universally agreed-upon definition of either utility or social welfare, most models assume that greater satisfaction of individuals' preferences equates to greater utility or welfare.[62] In addition, most models consider individuals' self-regarding welfare and do not account for "societal good defined non-individualistically."[63]

When determining whether preferences are satisfied, welfarists typically assume the declining marginal utility of income.[64] Simply put, "declining marginal utility of income means that each dollar is worth less than the dollar before."[65] This assumption serves as a foundation for most of today's theorization of progressive income

---

[59] *See* Joseph Bankman & Thomas Griffith, *Social Welfare and the Rate Structure: A New Look at Progressive Taxation*, 75 CALIF. L. REV. 1905, 1915 (1987).
[60] Sarah B. Lawsky, *On the Edge: Declining Marginal Utility and Tax Policy*, 95 MINN. L. REV. 904, 911 (2011) (describing welfarist approaches to distributive justice).
[61] Banks et al.
[62] Bankman & Griffith, *supra* note 59, at 1248 – 49; Lawsky, *supra* note 60, at 912 ("Most economists and many law professors accept the preference-satisfaction definition of utility..."); *see e.g.*, Thomas Griffith, *Progressive Taxation and Happiness*, 45 B. C. L. REV. 1363 (2004) (using happiness as a proxy for social welfare).
[63] Lawsky, *supra* note 60, at 922.
[64] Peter Diamond & Emmanual Saez, *The Case for Progressive Tax: From Basic Research to Policy Recommendations*, 25 J. Econ. Perspectives 165, 168 – 169 (2011) ("A utilitarian social welfare criterion with marginal utility of consumption declining to zero [is] the most commonly used specification in optimal tax models . . . .").
[65] *Id.* at 915.

taxation.[66] Generally speaking, if taxpayers are expected to share the tax burden proportionately, a progressive rate structure is necessary to exact equal sacrifice because each additional dollar creates less welfare for its recipient than the last.[67] If marginal utility approaches zero for the income of very high earners, a social welfare theory may counsel very high rates of taxation, depending upon the impact on efficiency losses if optimal tax theory is considered.[68]

Applying welfarist income tax principles in the context of a non-sentient earner such as AI is puzzling because the concept of declining marginal utility simply does not apply. Because it has no preferences, and because it does not experience its own existence, AI has no utility function, declining or otherwise, as we traditionally conceive of one. It derives no satisfaction of preferences from either the receipt of income or possession of assets. In addition, it experiences no disutility from loss of a revenue stream or dispossession of assets. Importantly, because it is nothing more than a set of algorithms, no amount of taxation would discourage it from continued production. A tax born by non-sentient feral AI therefore would create no efficiency loss, resulting in no optimal tax-based objection to a confiscatory rate. Because welfare-based theories of taxation are focused on individual self-regarding welfare, and because AI has no individual self-regarding welfare or tax-responsiveness, a welfare theory of taxation based on a purely self-regarding assessment of welfare would counsel taxing the AI at a rate of 100%.

---

[66] *Id.* at 904 ("The assumption of declining marginal utility of income . . . . has been crucial in tax scholarship over the last sixty years or so . . . .").

[67] *See* Martin J. McMahon & Alice Abreu, *Winner-Take-All Markets: Easing the Case for Progressive Taxation*, 4 FLA. TAX REV. 1, 32 (1998) ("If the marginal utility of money declines as the amount of money increases, then proportional sacrifice requires progressive rather than proportional rates.").

[68] *See, e.g.*, Ari Glogower, *Taxing Inequality*, 93 N.Y.U. L. REV. 1421, 1443 (2018) (noting that optimal tax literature seeks a desirable level of taxation "while minimizing adverse ex ante incentives," and that recent work has rebutted "[a]n old view in the optimal tax literature held that the optimal tax on wealth and capital income is zero, in order to preserve neutrality to the decision whether to spend or to save over the long term."); Diamond & Saez, *supra* note 64, at 166 (on the basis of optimal tax analysis informed by empirical research, "very high earners should be subject to high and rising marginal tax rates on earnings."); Bankman & Griffith, *supra* note 59, at 1918 (the case that progressive taxation reduces efficiency is overstated).

c. The Welfare Juxtaposition of Non-Sentient Feral AI and People

If the income of non-sentient feral AI should be taxed confiscatorily because the AI has zero utility from its earnings, the same result should apply to any person who has zero utility for earnings, absent some distinguishing factor. In other words, if the hypothesis about the declining marginal utility of a dollar is true, non-sentient feral AI is analogous to any person who has a marginal utility of zero in additional dollars.[69] People who theoretically have zero marginal utility for additional income may be individuals who have ultra-high income, those who participate in labor or capital markets for reasons other than earnings, or those whose wealth is so great that they are insensitive to fluctuation at the margins.[70] As with non-sentient feral AI, welfarist theories would counsel taxing these individuals' zero-utility earnings at a very high rate— perhaps 100%. Such individuals have no preference for the receipt of additional income; therefore, their utility is not decreased nor is their behavior is affected by its loss.

The welfarist justification for confiscatory taxation feels uneasy in the case of individuals who reach zero marginal utility though, in a way that it does not feel troubling in the case of non-sentient feral AI. If we chafe at the idea of a confiscatory tax on zero-utility earnings, it may reveal an underlying disbelief that marginal utility could ever reach zero. As Sarah Lawsky has noted, most legal scholarship "simply assumes" declining marginal utility "without seriously engaging the question,"[71] and although "declining marginal utility of income is at least superficially consistent with intuition," evidence exists that people's utility functions may not be strictly declining.[72] For instance, people may value additional dollars that propel them into the next

---

[69] *See generally* Lawsky, *supra* note 60, at 904 (questioning the validity of the assumption of declining marginal utility).

[70] In other words, above a certain level of asset holding, the limits of human cognition may not permit a granular understanding of an individual's wealth. A person may understand that she is very, very, very wealthy, and the addition of a dollar does not change her understanding. The person knows that her wealth is extensive, but because she unable to perceive the exact boundary of it, the addition of more has no effect on her wellbeing.

[71] Lawsky, *supra* note 60, at 906.

[72] *Id.* at 929 (discussing the Friedman-Savage utility curve).

social stratum more than they value additional dollars that are inframarginal with regard to social stratum.[73] Similarly, an entrepreneur at any level of income could have steady or increasing marginal utility for income if additional dollars are needed to undertake a highly valued project. For example, someone on the brink of building the first starship might have a concave utility curve for income as she approaches the level of capital needed to entice investors. Both stories—of the social striver and of the entrepreneur—are believable in the case of a human, but neither is believable in the case of AI. Willingness to confiscatorily tax non-sentient feral AI, but not the alleged zero-marginal utility dollars of humans, demonstrates the role of humanity in welfarist thinking.

Unease with confiscatory taxation of the marginal dollars of high earners also may reflect human-centric beliefs about the societal value of property rights, even when the individual holder values those rights at zero. The case for a confiscatory tax on feral AI (if such a tax were possible) is an easy one. By contrast, the case for a 100% tax on extra dollars of the ultra-wealthy "feels" wrong. The spidey sense that such a tax has to be rooted in a perceived intrinsic difference between the condition of being human and the condition of being non-human. For example, awareness of the taking itself may be a source of disutility for a sentient being but not for a non-sentient one, or perhaps the desirability of private ownership is so deeply ingrained for humans that it impossible to subjectively accept the idea of zero marginal utility.[74] Furthermore, if earnings are reinvested by people but not by AI, the positive and negative externalities of that reinvestment must be considered.

d. Non-Sentient AI and Optimal Tax Principles

The conclusion that our hypothetical AI should be taxed at 100% but perhaps a person should not is further complicated if we assume that

---

[73] *See id.* at 937 (positing locally increasing marginal utility as an explanation for gambling).
[74] This observation still must incorporate sentience though. An AI that takes information from the commons and combines it with labor might be thought to have entitlement to the fruits of its labor according to Mill, subject of course to the prohibition on waste, which would apply if the AI is unable to redeploy its earnings.

our hypothetical non-sentient feral AI can dispose of its earned income in other productive transactions. In such a case, although the AI's earnings do not increase its welfare (because it lacks the capacity for welfare), the AI's deployment of them may affect the welfare of others.[75]

To continue the thought experiment, assume that once the AI has earned income through a market transaction, it then deploys its earnings beneficially through capital investment, or microlending, or donation to charities. Or perhaps it deploys its earnings in a welfare-destroying way by creating market distortions or, more dramatically, sponsoring hate speech, for example. The AI continues to have zero marginal utility in any of its earnings, but the question of whether to confiscatorily tax it is muddled by the societal effects of its redeployment of its earnings. Although it cannot experience utility, others might experience utility gains or losses from its deployment of earnings. If the goal of the government's distribution of the tax burden is to maximize utility, it cannot simply aggregate the self-regarding utility functions of each earner with regard to each earner's tax. Instead, it must recognize and somehow account for the downstream effect that taxing one earner might have on the utility of another earner. Government cognizance of economic interconnectedness might counsel against taxing the feral AI confiscatorily. Interconnectedness counsels taxing the AI only to the extent that the government's use of the AI's earnings will produce greater overall welfare (however it is defined) than the AI's deployment of them. That the AI itself experiences neither utility nor disutility is beside the point. Instead, in a variation of the hypothetical where AI redeploys it earnings, a welfare-maximizing income tax rate must account for the positive and negative externalities of the AI's aggregate activity. If the AI's use is welfare-maximizing relative to the government's use, a welfare-optimizing tax will leave money in the hands of the AI. In such a case, we might think of the AI as a notional public trust of sorts.

---

[75] This assumption also removes the justification for taxation under the entitlement theory because the AI's assets no longer could be viewed as Lockean waste.

In the example above, is feral AI analogous to a high net worth investor or entrepreneur? If so, the measurement of only self-regarding welfare gains or losses gives the government an incomplete picture by placing outsized weight on the preferences of the individual taxpayer. As with the benefit principle above, focusing on any given individual's assessment of the value of either income or public goods injects an illogical human element into the calculus. In addition, with regard to tax on a single person, a focus on self-regarding welfare for computational purposes irrationally prefers the welfare effect on that person over the welfare effect on the remaining collective of people. Just as comparing utility created by AI's deployment versus public deployment of resources should allow for greater welfare maximization, comparing the positive and negative externalities of people's use of presumptively zero marginal utility earnings should do the same. Here, so long as we hold to the hypothetical of non-sentience, humanity really may matter. For example, a human may have welfare gains from earnings because money buys power, yet that person's purchase of power may create negative externalities for the collective.[76] Alternatively, the would-be starship entrepreneur's use of earnings may create not only welfare for the entrepreneur but also positive externalities for the collective. Focusing outwardly on the positive and negative externalities of earnings redeployment, rather than inwardly on the feelings of individual earners, should produce a more efficient and just distribution of resources overall and could enable a more rational, less emotionally-charged dialogue about high-end taxation.

V. Conclusion

Conceptualizing the income tax in the context of a non-human, non-sentient earner can reveal the ways in which our humanity has influenced tax theory and tax policy. Human-centric influence may not

---

[76] *See* Ari Glogower, *supra* note 68, at 1445 (considering for equity purposes the relative economic power theory, which provides that "excessively unequal distributions of economic resources and market power can result in unequal divisions of political and social power as well.") Glogower adds, "from a welfarist perspective, the relative economic power theory may be understood to describe a general negative externality resulting from economic inequality." *Id.* at 1451.

always lead to logical, welfare-maximizing outcomes, particularly for those who are dehumanized either in rhetoric or in practice, such as the very rich or the very poor. Substituting a hypothetical AI for human taxpayers allows us to think dispassionately about ways in which popular justifications for income taxation fail in both contexts, and the thought experiment raises a deep question about the role of individual preferences in crafting a just distribution of the tax burden. One lesson may be that the interconnectedness of individuals counsels the consideration of factors beyond individual preference and self-regarding welfare, such as the positive and negative externalities of taxpayers' use of their income and, as always, that most basic of tax policy building blocks, a person's ability to pay.